

FACTORING POLYNOMIALS OVER GLOBAL FIELDS

KARIM BELABAS, MARK VAN HOELJ¹, JÜRGEN KLÜNERS, AND ALLAN STEEL

ABSTRACT. Let K be a global field and $f \in K[X]$ be a polynomial. We present an efficient algorithm which factors f in polynomial time.

CONTENTS

1. Introduction	1
2. Notations	2
3. General description	3
3.1. Linearize	3
3.2. Approximately solve knapsack	3
3.3. Conclude	3
4. The case $K = \mathbb{Q}$	4
5. The case $K = \mathbb{F}_q(t)$	8
References	10

1. INTRODUCTION

Let K be a global field. The goal of this paper is to present a practical algorithm which factors polynomials $f \in K[X]$ in polynomial time, in particular for the cases $K = \mathbb{Q}$ and $K = \mathbb{F}_q(t)$. The seminal Zassenhaus [Zas69] method to factor in $K[X]$ is as follows: we may assume that f is separable, integral and monic. First, compute a bound for the factors of f , then find a non-archimedean place v of K such that the reduction \bar{f} of f modulo v remains separable in $k[X]$ where k is the residue field of v . Since k is finite we can factor \bar{f} in $k[X]$ using well known algorithms.

Let K_v the completion of K at v . Let \mathcal{O}_v , resp. \mathcal{O} be the maximal order of K_v , resp. K . If $K = \mathbb{Q}$ then v is a prime number, K_v the v -adic numbers, \mathcal{O}_v the v -adic integers, $\mathcal{O} = \mathbb{Z}$ and $k = \mathbb{Z}/v\mathbb{Z}$.

If $K = \mathbb{F}_q(t)$ then $\mathcal{O} = \mathbb{F}_q[t]$ and we will choose a finite place v , which corresponds to choosing an irreducible polynomial $v \in \mathbb{F}_q[t]$. If α is a root of this polynomial, then $k \cong \mathbb{F}_q(\alpha)$, $\mathcal{O}_v \cong k[[t - \alpha]]$ and $K_v \cong k((t - \alpha))$, see also Section 5.

After multiplying if necessary f by an element of \mathcal{O} , we may assume $f \in \mathcal{O}[X]$. By Hensel's lemma, the factorization of \bar{f} can be lifted to a factorization

$$f = \ell_f f_1 \cdots f_r$$

in $K_v[X]$ where $\ell_f \in K \subset K_v$ is the leading coefficient of f , and f_1, \dots, f_r are monic and irreducible in $K_v[X]$. We choose v so that $f \in \mathcal{O}_v[X]$ and ℓ_f does not vanish mod v , so $f_1, \dots, f_r \in \mathcal{O}_v[X]$. In actual computations, elements of \mathcal{O}_v are

¹Supported by NSF grant 0098034.

computed modulo v^ℓ for some $\ell > 0$ and lifted to \mathcal{O} . For $a \in \mathcal{O}_v$ we write “ $a \bmod v^\ell$ ” for such a lift of a to \mathcal{O} . This notation is extended to $\mathcal{O}_v[X]$ coefficientwise. By Hensel lifting the irreducible factors of \bar{f} we can compute $f_1, \dots, f_r \bmod v^\ell$ for any fixed $\ell > 0$.

Let $g \in K[X]$ be a monic irreducible factor of f . Then

$$g = f_1^{e_1} \cdots f_r^{e_r}$$

where $e_i \in \{0, 1\}$ for all $1 \leq i \leq r$. If ℓ is large enough compared to a bound on the coefficients of g , we may test for given $e_1, \dots, e_r \in \{0, 1\}$ whether $f_1^{e_1} \cdots f_r^{e_r} \in K[X]$ by computing $\ell_f f_1^{e_1} \cdots f_r^{e_r} \bmod v^\ell$ and checking whether this divides f in $K[X]$. This time, the lift “ $\dots \bmod v^\ell$ ” to $\mathcal{O}[X]$ is not arbitrary. Choosing the right lift is straightforward if K is \mathbb{Q} or $\mathbb{F}_q(t)$ since there are canonical minimal lifts to \mathcal{O} , but requires care for general global fields (see [Bel03] for the number field case).

The Zassenhaus algorithm finds the e_i by an exhaustive enumeration, which works very well if r is small or the K -rational irreducible factors are plentiful. Otherwise, we face combinatorial explosion and exponential behaviour. The landmark paper by Lenstra et al. [LLL82] avoids this combinatorial problem by constructing K -rational factors with lattice basis reduction (LLL reduction). The original paper assumes $K = \mathbb{Q}$, but was suitably generalized by Arjen Lenstra [Len82] (K a number field), then Pohst and Méndez [PO03] (K any global field). Unfortunately, although this algorithm runs in polynomial time, it is rather slow in practice since its worst case bounds require Hensel lift to huge accuracy, followed by the LLL-reduction of correspondingly huge lattices. Mark van Hoeij [Hoe02] came back to the combinatorial problem and used a knapsack approach to solve it for $K = \mathbb{Q}$, this was generalized to number fields by Belabas [Bel03]. These two papers stated no complexity bound. We shall describe a similar idea over a general global field K , and show it runs in polynomial time, although details will only be provided for the cases $K = \mathbb{Q}$ and $K = \mathbb{F}_q(t)$.

2. NOTATIONS

Throughout the paper we will use the following notations: Let K be a global field of characteristic $p \geq 0$ with maximal order \mathcal{O} . We want to factor a separable polynomial $f \in K[X]$ of degree $n > 1$. After multiplying by an element of K we may assume that $f \in \mathcal{O}[X]$. Let v be a non-archimedean place of \mathcal{O} . We denote by K_v the completion of K at v , with maximal order \mathcal{O}_v , maximal ideal v and finite residue field k . Let \bar{f} be the image of f in $k[X]$, and assume that \bar{f} is still separable. We also assume that the leading coefficient $\ell_f \in \mathcal{O}$ of f does not vanish mod v , so the degree of \bar{f} is still n . In the number field case, instead of working with \mathcal{O} we can work with a subring of \mathcal{O} if the computation of \mathcal{O} is too costly, see [Bel03].

We have the factorizations into irreducible elements

$$f = \ell_f f_1 \cdots f_r \in \mathcal{O}_v[X], \quad \bar{f} = \bar{\ell}_f \bar{f}_1 \cdots \bar{f}_r \in k[X], \quad \text{and} \quad f = \ell_f g_1 \cdots g_s \in K[X].$$

Furthermore, $\ell_f g_i \in \mathcal{O}[X]$. Obviously $1 \leq s \leq r \leq n$. We call the f_i the *local factors* and the g_j the *K -factors*. We can not compute $f_i \in \mathcal{O}_v[X]$ with infinite accuracy, but for any positive integer ℓ we can compute $f_i \bmod v^\ell$, which is in $\mathcal{O}[X]$.

3. GENERAL DESCRIPTION

Our method relies on two main ideas:

3.1. Linearize. The logarithmic derivative is a group homomorphism from the multiplicative group $K_v(X)^*$ to the additive group $K_v(X)$, and has kernel $K_v(X^p)^*$. The first main idea is to multiply this by f . Then we obtain the following group homomorphism:

$$\begin{aligned} \Phi : K_v(X)^*/K_v(X^p)^* &\rightarrow K_v(X) \\ g &\mapsto fg'/g. \end{aligned}$$

If g is in the subgroup of $K_v(X)^*/K_v(X^p)^*$ generated by the local factors f_i , then $\Phi(g) \in \mathcal{O}_v[X]$. If g is in the subgroup generated by the K -factors g_j , then $\Phi(g) \in \mathcal{O}[X]$. To see this, take one such g_j . Take any prime ideal of \mathcal{O} and let w be the corresponding valuation on K , which is extended to a valuation on $K[X]$ by taking $w(\sum c_i X^i) = \min_i w(c_i)$. Now $\Phi(g_j)$ is the product of g'_j and f/g_j , both of which are in $K[X]$, but since $w(g'_j) \geq w(g_j)$ we get $w(fg'_j/g_j) \geq w(fg_j/g_j)$ which is ≥ 0 since $f \in \mathcal{O}[X]$. So the valuation of $\Phi(g_j) \in K[X]$ is non-negative for any prime ideal of \mathcal{O} and hence $\Phi(g_j) \in \mathcal{O}[X]$.

Compared to the original algorithm of van Hoeij [Hoe02], we have replaced power sums by f times the logarithmic derivative. To show the connection we will define power sums. Let $g \in K[X]$ be a monic separable polynomial. Let $\alpha_1, \dots, \alpha_m$ the zeros of g in an algebraic closure of K . For $j \geq 0$, the j 'th power sum (j -th “trace”) of g is:

$$\text{Tr}_j(g) := \sum_{i=1}^m \alpha_i^j.$$

It is known that

$$g'/g = \sum_{j \geq 0} \text{Tr}_j(g) X^{-j-1},$$

which shows the relation between g'/g and power sums. Despite this relation, our “ f times g'/g approach” turns out to be more convenient for complexity proofs than power sums, and can also have practical advantages, particularly when f is not monic.

3.2. Approximately solve knapsack. Let $G_v \subset K_v(X)^*/K_v(X^p)^*$ be the subgroup generated by the local factors. Our goal is to find the subgroup $G \subset G_v$ generated by the irreducible K -factors of f . To do this we first construct the “knapsack lattice” L in a similar way as in [Hoe02], except that instead of traces (power sums) of f_j we use the coefficients of $\Phi(f_j)$. We then reduce this lattice, which means \mathbb{F}_p -linear Gaussian elimination if $p > 0$, and LLL otherwise, for details see Sections 4 and 5. Large basis vectors are then discarded, yielding a sublattice L' of L , associated to a subgroup G' of G_v .

3.3. Conclude.

Theorem 3.1. *We have $G = G'$ provided ℓ is large enough.*

Proof. We will only sketch the proof, leaving the details to Sections 4 and 5. If G' is strictly larger than G , then by Lemma 3.2 below, it contains an element, represented by a rational function $g \in K_v(X)^*$, $g \notin K_v(X^p)^*$, such that

- (1) At least one f_i divides $\Phi(g)$,

- (2) None of the $\overline{g_j}$ divide $\overline{\Phi(g)}$ where the bar indicates reduction to $k[X]$.
- (3) $H := \Phi(g) \bmod v^\ell$ is “small”.

Indeed, it is a small perturbation (see Lemma 3.2 below) of a vector in the LLL basis of L' that is small otherwise it would have been discarded.

The clumsy argument for the third condition is only needed if $p = 0$. If $p > 0$, all elements of L' are small. Now, let g be as above, $H := \Phi(g) \bmod v^\ell$, and $R := \text{Res}(f, H)$. Then

- $\text{Res}(f, \Phi(g)) = 0$, hence $v^\ell \mid R$. In fact, $v^{\ell\sigma} \mid R$ where σ is the sum of the degrees of the f_i that divide $\Phi(g)$. Item (1) above implies $\sigma > 0$.
- $R \neq 0$, because if R was zero then H would be divisible by some g_j so \overline{H} would be divisible by some $\overline{g_j}$, contradicting item (2) above.

One obtains a contradiction if v^ℓ is larger than R . Since R is the determinant of the Sylvester matrix of f and H , one obtains a bound that is polynomial in terms of the sizes of f and H . \square

Lemma 3.2. *Suppose $G \subsetneq G'$. Then there exists an element $g \in G' \setminus G$ such that*

- (1) $f_i \mid \Phi(g) \in \mathcal{O}_v[X]$ for some $1 \leq i \leq r$.
- (2) $\overline{g_j} \nmid \overline{\Phi(g)}$ for all $1 \leq j \leq s$.

Proof. Elements $g \in G_v$ can be written in the form $g = f_1^{e_1} \cdots f_r^{e_r} \cdot K_v(X^p)^*$ where the integers e_i are defined mod p . We view e_i as element of $\mathbb{Z}/p\mathbb{Z}$, and then define the support of g as $\text{Supp } g = \{i \mid e_i \neq 0\}$. Since the f_i are pairwise coprime and irreducible in $\mathcal{O}_v[X]$, we have

$$f_i \mid \Phi(f_j) \iff i \neq j.$$

So $f_i \mid \Phi(g)$ iff e_i is zero in $\mathbb{Z}/p\mathbb{Z}$, and $g_j \mid \Phi(g)$ iff $\text{Supp } g \cap \text{Supp } g_j = \emptyset$.

The supports of g_1, \dots, g_s form a partition of $\{1, \dots, r\}$. Choose any element $g \in G' \setminus G$. For all $1 \leq j \leq s$ with $\text{Supp } g_j \cap \text{Supp } g = \emptyset$, replace g by $g_j g$. Then condition (2) is satisfied (recall that \overline{f} is separable), and g is still in $G' \setminus G$. Write this g as $f_1^{e_1} \cdots f_r^{e_r} \cdot K_v(X^p)^*$ with $e_i \in \mathbb{Z}/p\mathbb{Z}$. Since g is not in the group G generated by g_1, \dots, g_s , there must be some g_j for which $S_j := \{e_i \mid i \in \text{Supp } g_j\}$ contains more than one element. Then take an element $e \in S_j$ and replace g by g/g_j^e . Now g satisfies both conditions (1) and (2). \square

Remark 3.3. Given any $g \in G' \setminus G$, the above proof shows that a “small change” suffices to obtain an element of $G' \setminus G$ that satisfies conditions (1) and (2).

We have sketched a general proof and omitted the details. Filling in these details is easy for the case $K = \mathbb{F}_q(t)$ discussed in Section 5. The details for $K = \mathbb{Q}$ require more work, which is what we will do now.

4. THE CASE $K = \mathbb{Q}$

For $f \in \mathbb{C}[X]$ with leading coefficient ℓ_f , let

$$M(f) := |\ell_f| \prod_{|\alpha| > 1} |\alpha|^{m_\alpha}$$

be the Mahler measure of f , where the product is taken over all roots $\alpha \in \mathbb{C}$ of f with absolute value > 1 , and m_α is the multiplicity of the root α .

Lemma 4.1. *If $f, g \in \mathbb{C}[X]$ and $g \mid f$ then*

$$\Phi(g) = \sum_{i=0}^{n-1} a_i X^i \in \mathbb{C}[X], \text{ with } |a_i| \leq B_i := \binom{n-1}{i} nM(f).$$

Proof. We may assume g is not a constant. Then the degree of $\Phi(g) \in \mathbb{C}[X]$ is $n-1$. The Mahler measure of $\Phi(g)$ is bounded by $\deg(g)M(f)$ since $M(A') \leq \deg(A)M(A)$, see [Mah61], and $M(AB) = M(A)M(B)$ for any $A, B \in \mathbb{C}[X]$ [MS99, p. 79]. Bounding $\deg(g)$ by n , the upper bound now follows by [MS99, Lemma 2.1.9]. \square

We restrict to the case $K = \mathbb{Q}$, for a general number field follow [Bel03]. We use the notation $\mathbb{Z}[X]_{<n}$ for all polynomials in $\mathbb{Z}[X]$ of degree $< n$. We use $\|\cdot\|_2$ for the L^2 norm on both \mathbb{Z}^n and $\mathbb{Z}[X]_{<n}$.

Corollary 4.2. *With $f \in \mathbb{Z}[X]$ and g any factor of f in $\mathbb{Q}[X]$, we have $\Phi(g) \in \mathbb{Z}[X]_{<n}$ and*

$$\|\Phi(g)\|_2 \leq B(f) := 2^{n-1}n \|f\|_2.$$

Proof. That $\Phi(g)$ is in $\mathcal{O}[X]$ was proven in Section 3.1. Using Lemmata 2.1.8 and 2.1.9 in [MS99] we get that $\|\Phi(g)\|_2 \leq 2^{n-1}M(\Phi(g))$. As in the proof of Lemma 4.1 we get that $2^{n-1}M(\Phi(g)) \leq 2^{n-1}nM(f)$. Corollary 2.1.5 in [MS99] states that $M(h) \leq \|h\|_2$ for all non constant polynomials h which finishes the proof. \square

For $1 \leq j \leq s$ write the monic irreducible K -factors as $g_j = f_1^{w_{j,1}} \cdots f_r^{w_{j,r}}$ with $w_{j,1}, \dots, w_{j,r} \in \{0, 1\}$ and write $w_j := (w_{j,1}, \dots, w_{j,r})^{\text{tr}} \in \mathbb{Z}^r$ where tr denotes the transpose. Denote $W = \mathbb{Z}w_1 + \cdots + \mathbb{Z}w_s$.

If we have any basis u_1, \dots, u_s of W then we can find $\{w_1, \dots, w_s\}$ by computing the reduced echelon form of u_1, \dots, u_s , or by using the following shortcut: write $\{1, \dots, r\}$ as a disjoint union of subsets in such a way that i, j are in the same subset iff the i 'th and j 'th entry of u are the same for every u in u_1, \dots, u_s .

In the following let I_r be the identity matrix of dimension r and define for $1 \leq j \leq r$ the $a_{i,j}$ via

$$\Phi(f_j) \bmod v^\ell = \sum_{i=0}^{n-1} a_{i,j} x^i.$$

Define the *all-coefficients* lattice L as the span of the columns of the following matrix:

$$A := \begin{pmatrix} I_r \\ A_1 \end{pmatrix}, \text{ where } A_1 := \begin{pmatrix} a_{0,1} & \cdots & a_{0,r} \\ \vdots & \ddots & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,r} \end{pmatrix}.$$

For $e = (e_1, \dots, e_{r+n})^{\text{tr}} \in L$ we denote the corresponding element $\Phi(f_1^{e_1} \cdots f_r^{e_r})$ of $\Phi(G_v)$ as $\text{POL}(e)$. Each K -factor g_j corresponds to a vector in L we denote by \tilde{w}_j , whose entries come from w_j and $\Phi(g_j)$. Then $\|\tilde{w}_j\|_2 \leq \sqrt{\|w_j\|_2^2 + B^2} \leq B' := \sqrt{r^2 + B^2}$ where $B = B(f)$ is as in Corollary 4.2.

Theorem 4.3. *Let $f \in \mathbb{Z}[X]$ separable and B' as above. Let b_1, \dots, b_n an LLL-reduced basis for the all-coefficients lattice L , let b_1^*, \dots, b_n^* the associated Gram-Schmidt orthogonalized basis, and let t the smallest index such that $\|b_j^*\|_2 > B'$ for all $j > t$. Let $L' := \mathbb{Z}b_1 + \cdots + \mathbb{Z}b_t$. If*

$$(1) \quad v^{\ell/n} > \|f\|_2 (2^{n-1} + n)B'(1 + B').$$

then the projection of L' on the first r entries is W .

Proof. It follows from the proof of (1.11) in [LLL82] that every $w \in L$ with $\|w\|_2 \leq B'$ is in L' . So $\tilde{w}_1, \dots, \tilde{w}_s \in L'$ and hence the projection of L' on the first r entries contains W . Assume it is strictly larger, then $\text{POL}(b_u) \notin \Phi(G)$ for some $1 \leq u \leq t$. From the properties of LLL-reduced bases, $\|b_u\|_2 \leq \gamma^{t-1} B'$, where $\gamma > 4/3$ is a number that can be chosen in the reduction algorithm (one may set $\gamma := 2$ as in the original LLL paper). More precisely, $\|b_u\|_2 \leq \gamma^{t-u} \|b_t^*\|_2$ and $\|b_u\|_2 \leq \gamma^{u-1} \|b_u^*\|_2$ for all $u \leq t \leq n$.

Using Lemma 3.2 as in the proof of Theorem 3.1 one can show that there exists a vector $g \in L'$ such that $f_i \mid \text{POL}(g)$ for some $1 \leq i \leq r$, and $v^\ell \mid \text{Res}(f, H) \neq 0$, where $H := \text{POL}(g) \bmod v^\ell$. From the proof of Lemma 3.2, this vector g may be obtained by first adding a subset of $\{\tilde{w}_1, \dots, \tilde{w}_s\}$ to b_u , yielding a vector b such that

$$\|b\|_2 \leq (\gamma^{t-1} + s)B',$$

and then by adding to b a vector of the form $e\tilde{w}_i$ for some integer e with $|e| \leq \|b\|_\infty \leq \|b\|_2$. Hence

$$\|H\|_2 \leq \|g\|_2 \leq (\gamma^{t-1} + s)B'(1 + B').$$

From the preceding discussion and Hadamard's bound,

$$v^\ell \leq |\text{Res}(f, H)| \leq \|f\|_2^{\deg H} \|H\|_2^n,$$

and we may bound $\deg H, s, t \leq n$ to derive a contradiction with (1). \square

From this theorem, we obtain $\ell \log v = O(n^2 + n \log \|f\|_2)$. Since Hensel lifting and lattice reduction are polynomial time algorithms, we see that W can be computed in polynomial time.

Although there is no practical reason for doing so (since power sums do not offer advantages over coefficients of Φ), one could now use the relation between power sums (called traces in [Hoe02]) and Φ to show the algorithm in [Hoe02] is polynomial time provided that one uses what we call the *all-traces* version of the algorithm. This version uses all of the traces numbered $1, \dots, n-1$ at the same time, so the lattice reduction takes place in \mathbb{Z}^{r+n-1} . From a practical point of view, the all-traces and all-coefficients versions are slow and thus not interesting.

The main question is whether practical versions of the algorithm run in polynomial time. Using one trace at a time works very well in practice, see [Bel03]. We will show that the “one coefficient at a time” version factors in $\mathbb{Q}[X]$ in polynomial time (the same must then also be true for one trace at a time).

Let B_i be the bound for $|a_i|$ given in Lemma 4.1. For $0 \leq i \leq n-1$ and $g \in K_v(X)^*$ write $T'_i(g) \in \mathbb{Z}$ the coefficient of X^i in $\Phi(g) \bmod v^\ell$. Let $T_i(g) := T'_i(g)/B_i \in \mathbb{Q}$. Now Lemma 4.1 says that if g is a K -factor of f , then $|T_i(g)| \leq 1$.

Proposition 4.4. *One can compute a sequence of lattices $L_{n-1}, L_{n-2}, \dots, L_0$ with the following properties:*

- (1) $\mathbb{Z}^r = L_{n-1} \supseteq L_{n-2} \cdots \supseteq L_0 \supseteq W$
- (2) $L_i = \mathbb{Z}b_{i,1} + \cdots + \mathbb{Z}b_{i,r_i}$ for some integer r_i and some vectors $b_{i,j} \in \mathbb{Z}^r$ with the following properties:
 - (a) $\|b_{i,j}\|_2 \leq (r+2)\gamma^r$.
 - (b) If $b_{i,j} = (e_1, \dots, e_r)^{\text{tr}}$ then $T_i(f_1^{e_1} \cdots f_r^{e_r}) \leq (r+2)\gamma^r$

where $\gamma > 4/3$ is a number that can be chosen in the reduction algorithm (one may set $\gamma := 2$ as in the original LLL paper).

Proof. If $i = n - 1$ we may take $b_{i,1}, \dots, b_{i,r_i}$ as the standard basis of \mathbb{Z}^r . If $i < n - 1$ then we may assume that $L_{i+1} = \mathbb{Z}b_{i+1,1} + \dots + \mathbb{Z}b_{i+1,r_{i+1}}$ has been computed and define b'_j as follows: First write $b_{i+1,j} = (e_1, \dots, e_r)^{\text{tr}}$, then compute $a := e_1 T_i(f_1) + \dots + e_r T_i(f_r)$ and set $b'_j := (e_1, \dots, e_r, a)^{\text{tr}} \in \mathbb{Z}^r \times \mathbb{Q}$. Now let $L' := \mathbb{Z}b'_1 + \dots + \mathbb{Z}b'_{r_{i+1}} + \mathbb{Z}P$ where $P = (0, \dots, 0, v^\ell/B_i)^{\text{tr}}$. Let b_1, b_2, \dots be an LLL-reduced basis of L' , let b_1^*, b_2^*, \dots the associated orthogonalized basis, and let r_i be the smallest index such that $\|b_j^*\|_2 > r + 2$ for all $j > r_i$. Now define $b_{i,j}$ as the projection of b_j on the first r entries and let $L_i := \mathbb{Z}b_{i,1} + \dots + \mathbb{Z}b_{i,r_i}$.

Consider the vector w_j corresponding to the K -factor g_j and let w'_j be the corresponding vector in L' . The first r entries of w'_j are in $\{0, 1\}$, and the last entry equals $T_i(g_j) \in \mathbb{Q}$ which has absolute value ≤ 1 by Lemma 4.1. Hence, $\|w'_j\|_2 \leq \sqrt{r+1} < r+2$. Then it follows from the proof of (1.11) in [LLL82] that $w_j \in L_i$ and hence $W \subseteq L_i$. By the properties of an LLL-reduced basis, we have $\|b_j\|_2 \leq (r+2)\gamma^r$ when $j \leq r_i$ which implies (2a) resp. (2b) since projecting on the first r entries resp. last entry does not make a vector longer.

The lattice L' to be reduced was in $\mathbb{Z}^r \times \mathbb{Q}$. Lattice reduction in \mathbb{Z}^{r+1} is more efficient, so we round each of the numbers $T_i(f_1), \dots, T_i(f_r), v^\ell/B_i$ to the nearest integer. Then we obtain a lattice $L' \subseteq \mathbb{Z}^{r+1}$ but now we have introduced rounding errors. Consider again the vectors $w_j \in W$ and $w'_j \in L'$. If w_j has σ entries equal to 1, then the last entry of w'_j is the sum of σ of elements of $\{T_i(f_1), \dots, T_i(f_r)\}$ plus an integer in the interval $(-\sigma/2, \sigma/2)$ times v^ℓ/B_i . We introduced an error ≤ 0.5 in each of the numbers $T_i(f_1), \dots, T_i(f_r), v^\ell/B_i$. Then the total rounding error in the last entry of w'_j is less than $0.5(\sigma + \sigma/2)$ which is less than r , so this entry will have absolute value $< r + 1$. Then $\|w'_j\|_2 < \sqrt{\sigma + (r+1)^2} < r + 2$. The proposition is stated with $r + 2$ instead of $\sqrt{r+1}$ so that the bounds can still be used for practical implementations that round $T_i(f_1), \dots, T_i(f_r), v^\ell/B_i$ to \mathbb{Z} . \square

If L_{i+1} is known, then the computation of L_i in the proposition involves a lattice reduction in \mathbb{Z}^{r+1} of a lattice with determinant v^ℓ/B_i (rounded to the nearest integer). If v^ℓ/B_i is large, then we get a big practical improvement by doing this lattice reduction incrementally in the way it is described in Section 2.4 in [Bel03], reducing one large-determinant lattice reduction to a sequence of smaller lattice reductions that at the end produce the same result.

Lemma 4.5. *With the notations of Proposition 4.4, the following holds for every $n - 1 \geq i \geq i' \geq 0$. If $e = (e_1, \dots, e_r)^{\text{tr}}$ is an element of $\{b_{i',1}, \dots, b_{i',r_{i'}}\}$ then*

$$T_i(f_1^{e_1} \dots f_r^{e_r}) \leq 2^{O(r^2)}$$

Proof. The entries of the $b_{i,j}$ and e are bounded by $(r+2)\gamma^r = 2^{O(r)}$. Since $e \in L_{i'} \subseteq L_i$ we can write $e = \sum_{j=1}^{r_i} c_j b_{i,j}$ for some $c_j \in \mathbb{Z}$ that can be found by solving linear equations. With Cramer's rule one finds $|c_j| \leq 2^{O(r^2)}$. Multiplying this by r_i and by the bound given in (2b) in Proposition 4.4 one obtains the bound $2^{O(r^2)}$. \square

Theorem 4.6. *$L_0 = W$ for some ℓ with $\ell \log v$ polynomially bounded in terms of the degree of f and $\log \|f\|_2$.*

Proof. If $L_0 \neq W$ then let e be one of the vectors $b_{0,j}$ from Proposition 4.4 that is not in W . Write $e = (e_1, \dots, e_r)^{\text{tr}}$ and $g = f_1^{e_1} \cdots f_r^{e_r}$. Write $\Phi(g) = \sum c_i X^i$. Then the corresponding vector in the all-coefficients lattice (see Theorem 4.3) is $\tilde{e} := (e_1, \dots, e_r, c_0, \dots, c_{n-1})^{\text{tr}}$ where c_0, \dots, c_{n-1} are bounded in absolute value by $2^{O(r^2)}$ by Lemma 4.5. Applying the process in the proof of Lemma 3.2 we obtain a new vector e' whose length differs at most $(s + \max\{e_1, \dots, e_r\})B'$ from e . The last n entries of this vector are the coefficients of a polynomial $H \in \mathbb{Z}[X]_{<n}$ and we have $v^\ell \mid \text{Res}(f, H) \neq 0$ in the same way as in Theorem 4.3. This implies that $\log v^\ell$ is polynomially bounded. \square

We propose to implement the “one coefficient at a time” approach in the following way: start with a value for ℓ that is *at most* as large as what one would use in the Zassenhaus approach. Then, compute L_{n-1}, L_{n-2}, \dots until we find W . If we reach L_0 and we still have not found W then we must increase ℓ . The computation of each L_i should be done using the incremental strategy of Section 2.4 in [Bel03]. Then one has a polynomial time algorithm that runs very well in practice, with running times that are essentially the same as those reported in [Bel03] for $\mathbb{Q}[X]$.

5. THE CASE $K = \mathbb{F}_q(t)$

Now $\mathcal{O} = \mathbb{F}_q[t]$, and the place v corresponds to an irreducible polynomial in $\mathbb{F}_q[t]$, which we shall also denote as v . Let $f \in \mathcal{O}[X]$. We want to factor f , viewed as element of $\mathbb{F}_q(t)[X]$. We assume that f is separable. Denote α as a root of $v \in \mathbb{F}_q[t]$, then the residue field $k = \mathbb{F}_q[t]/(v)$ is isomorphic to $\mathbb{F}_q(\alpha)$. We choose v in such a way that \bar{f} , the image of f in $k[X]$, is squarefree and of the same X -degree as f . We get the factorization

$$\bar{f} = \bar{\ell}_f \bar{f}_1 \cdots \bar{f}_r \in k[X].$$

Representing $t - \alpha$ with a new variable \tilde{t} , the map $t \mapsto \tilde{t} + \alpha$ is an isomorphism from $\mathbb{F}_q[t]/(v^\ell)$ to $\mathbb{F}_q(\alpha)[\tilde{t}]/(\tilde{t}^\ell)$. Taking limits, one finds an isomorphism from

$$\mathcal{O}_v = \varprojlim \mathbb{F}_q[t]/(v^\ell)$$

to

$$\mathbb{F}_q(\alpha)[[\tilde{t}]] = \varprojlim \mathbb{F}_q(\alpha)[\tilde{t}]/(\tilde{t}^\ell).$$

By Hensel’s lemma, we get a factorization

$$f = \ell_f f_1 \cdots f_r \in \mathcal{O}_v[X].$$

If $g \in \mathcal{O}_v[X]$ we denote “ $g \bmod v^\ell$ ” as the unique lift of g to $\mathbb{F}_q[t, X]$ whose t -degree is smaller than the t -degree of v^ℓ . We can not compute $f_i \in \mathcal{O}_v$ with infinite accuracy, however, for any integer $\ell > 0$ we can compute $f_i \bmod v^\ell$, which is an element of $\mathbb{F}_q[t, X]$.

Note that the above technicalities with \mathcal{O}_v become easier if we take $v = t$ so that $\tilde{t} = t$. However, we can not always do this, we can only take $v = t$ if $f(t = 0, X)$ is square-free and of the same degree as f .

Lemma 5.1. *Let $g \in \mathbb{F}_q[t][X]$ be a polynomial which divides f then*

$$\Phi(g) = \sum_{i=0}^{n-1} a_i(t) x^i \in \mathbb{F}_q[t][x] \text{ with } \deg(a_i) \leq B_i := \deg_t(f),$$

where \deg_t denotes the t -degree.

Proof. From $\Phi(g) = fg'/g$ we get $\deg_t(\Phi(g)) + \deg_t(g) = \deg_t(g') + \deg_t(f)$. Since $\deg_t(g') \leq \deg_t(g)$ we get the wanted bound. \square

The idea is as follows. Let $g \in G_v$. If the degree of one of the coefficients of $\Phi(g) \bmod v^\ell$ exceeds the degree bound B_i then g is not a K -factor of f . We use this to replace the Zassenhaus combinatorial search by linear algebra.

As in the rational case we introduce the lattice $W \subseteq (\mathbb{Z}/p\mathbb{Z})^r$ generated by the exponent vectors of the monic irreducible factors g_1, \dots, g_s of f in $\mathbb{F}_q(t)[X]$. Let L be some subspace of $(\mathbb{Z}/p\mathbb{Z})^r$ that contains W . We start with $L = (\mathbb{Z}/p\mathbb{Z})^r$. For $e = (e_1, \dots, e_r)^{\text{tr}} \in L$ we denote by $\text{POL}(e)$ the polynomial

$$\Phi(f_1^{e_1} \cdots f_r^{e_r}) \bmod v^\ell.$$

Our goal is to compute a subspace of $L' \subseteq L$ which still contains W . Write

$$\text{POL}(\varepsilon_j) = \sum_{i=0}^{n-1} a_{i,j} X^i \quad (1 \leq j \leq r)$$

where $\varepsilon_1, \dots, \varepsilon_r$ is the standard basis of $(\mathbb{Z}/p\mathbb{Z})^r$. Let $m_i = B_i + 1$ and let σ be the t -degree of v^ℓ . We define

$$\phi_{m_i} \left(\sum_k c_k t^k \right) := (c_{m_i}, \dots, c_{\sigma-1})^{\text{tr}}$$

and

$$A_i := (\phi_{m_i}(a_{i,1}) \cdots \phi_{m_i}(a_{i,r}))$$

which is an $(\sigma - m_i) \times r$ matrix with entries in \mathbb{F}_q , and $A_i e = 0$ for all $e \in W$. Now L and W are subspaces of \mathbb{F}_p^r and A_i is defined over \mathbb{F}_q . For $q = p^w$ write $\mathbb{F}_q = \mathbb{F}_p \gamma_1 + \cdots + \mathbb{F}_p \gamma_w$ and define

$$\psi : \mathbb{F}_q \rightarrow \mathbb{F}_p^w, \quad \sum_{l=1}^w c_l \gamma_l \mapsto (c_1, \dots, c_w)^{\text{tr}},$$

where tr denotes the transpose. We define \tilde{A}_i as follows: replace every entry c of A_i by $\psi(c)$. Since $\psi(c)$ is a column vector (because of the transpose in its definition) with w entries we see that \tilde{A}_i is an $w(\sigma - m_i) \times r$ matrix with entries in \mathbb{F}_p . We still have $\tilde{A}_i e = 0$ for all $e \in W$. Now let L' be the intersection of the kernels of $\tilde{A}_0, \dots, \tilde{A}_{n-1}$. Then L' contains W .

Let B be a degree bound which can be easily computed using Theorem 3.1. E.g. we can take $B = (2n - 1) \deg_t(f)$ when we use the estimate from Lemma 5.1 and the properties of the Sylvester matrix. Theorem 3.1 guarantees that L' will be W when σ (the t -degree of v^ℓ) is larger than B . Altogether we have proved

Theorem 5.2. *If the t -degree of v^ℓ is larger than $(2n - 1) \deg_t(f)$ and L' is the intersection of the kernels of \tilde{A}_i , $i = 0, \dots, n - 1$ then $L' = W$. This leads to an algorithm that produces the factorization of a separable polynomial $f \in \mathbb{F}_q[t][X]$ in polynomial time.*

Remark 5.3. If the total degree of f as bivariate polynomial is n , then one can replace $B_i := \deg_t(f)$ in Lemma 5.1 by $B_i := n - 1 - i$. The proof is essentially the same, except that the degree w.r.t. t should be replaced by the total degree. Then we can replace $(2n - 1) \deg_t(f)$ by $n(n - 1)$ in the above theorem.

Remark 5.4. In [BLSSW04] the authors followed our “ f times g'/g ” approach found in a previous version of this paper and were able to improve the quadratic bound $n(n-1)$ to a linear bound when $p > n(n-1)$.

Note that in an implementation, one would start with a small value for ℓ , increasing ℓ as long as L' is not W . To improve practical performance, we can replace the bounds B_i from Lemma 5.1 or Remark 5.3 by the sharper bound given in the lemma below.

Denote $N(f) \subset \mathbb{R}^2$ as the Newton polygon of f , which is defined as the convex hull of all points (i, j) for which the coefficient of $t^i X^j$ in f is non-zero. If $S_1, S_2 \subset \mathbb{R}^2$ then define $S_1 + S_2 := \{s_1 + s_2 \mid s_1 \in S_1, s_2 \in S_2\}$.

Lemma 5.5. *Let $B_i := \sup\{m \in \mathbb{N} \mid (m, i) \in N(f) + \{(0, -1)\}\}$. Let $g \in \mathbb{F}_q[t][X]$ be a polynomial which divides f then*

$$\Phi(g) = \sum_{i=0}^{n-1} a_i(t)x^i \in \mathbb{F}_q[t][x] \text{ with } \deg(a_i) \leq B_i.$$

Proof. It is well known that $N(gh) = N(g) + N(h)$ for all $g, h \in \mathbb{F}_q[t, X]$. It is also clear that $N(g') \subseteq N(g) + \{(0, -1)\}$. Then $N(\Phi(g)) = N(f/g \cdot g') = N(f/g) + N(g') \subseteq N(f/g) + N(g) + \{(0, -1)\} = N(f) + \{(0, -1)\}$. \square

REFERENCES

- [Bel03] K. Belabas *A relative van Hoeij algorithm over number fields*, Journal of Symbolic Computation, to appear. Cf. <http://www.math.u-psud.fr/~belabas/pub/#vanhoeij>
- [BLSSW04] A. Bostan, G. Lecerf, B. Salvy, E. Schost, and B. Wiebelt. *Complexity issues in bivariate polynomial factorization* Proceedings of ISSAC, 2004
- [Hoe02] M. van Hoeij, *Factoring polynomials and the knapsack problem*, J. Number Theory, **95**, (2002), 167–189.
- [Len82] A. K. LENSTRA, Lattices and factorization of polynomials over algebraic number fields, (Berlin), LNCS, vol. 144, Springer, Berlin, 1982, pp. 32–39.
- [LLL82] A. K. LENSTRA, H. W. LENSTRA, JR., AND L. LOVÁSZ, Factoring polynomials with rational coefficients, *Math. Ann.* **261** (1982), no. 4, pp. 515–534.
- [MS99] D. Stefanescu and M. Mignotte, *Polynomials*, Springer, 1999.
- [Mah61] K. Mahler, *Proc. Roy. Soc. Ser. A*, **264**, (1961), 145–154.
- [PZ89] M. E. Pohst and H. Zassenhaus, *Algorithmic algebraic number theory*, Cambridge University Press, 1989.
- [PO03] M. E. Pohst and J. Méndez, *Factoring polynomials over global fields*, preprint.
- [Zas69] H. ZASSENHAUS, On Hensel factorization I, *Journal of Number Theory* (1969), pp. 291–311.

UNIVERSITÉ PARIS-SUD, DÉPARTEMENT DE MATHÉMATIQUE, 91400 ORSAY, FRANCE.

E-mail address: Karim.Belabas@math.u-psud.fr

FLORIDA STATE UNIVERSITY, 211 LOVE BUILDING, TALLAHASSEE, FLORIDA 32306-3027, USA

E-mail address: hoeij@zeno.math.fsu.edu

UNIVERSITÄT KASSEL, FACHBEREICH MATHEMATIK UND INFORMATIK, HEINRICH-PLETT-STR. 40, 34132 KASSEL, GERMANY.

E-mail address: klueners@mathematik.uni-kassel.de

SCHOOL OF MATHEMATICS AND STATISTICS F07, UNIVERSITY OF SYDNEY NSW 2006, AUSTRALIA

E-mail address: allan@maths.usyd.edu.au